

Research Article

Assessing Traffic Accident Occurrence of Road Segments through an Optimized Decision Rule

Lu Ma¹ and Xuedong Yan²

¹MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

²Center of Cooperative Innovation for Beijing Metropolitan Transportation, School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Xuedong Yan; xdyan@bjtu.edu.cn

Received 12 July 2014; Accepted 28 October 2014

Academic Editor: Huimin Niu

Copyright © 2015 L. Ma and X. Yan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Statistical models for estimating the safety status of transportation facilities have received great attention in the last two decades. These models also perform an important role in transportation safety planning as well as diagnoses of locations with high accident risks. However, the current methods largely rely on regression analyses and therefore they could ignore the multicollinearity characteristics of factors, which may provide additional information for enhancing the performance of forecasting models. This study seeks to develop more precise models for forecasting safety status as well as addressing the issue of multicollinearity of dataset. The proposed mathematical approach is indeed a discriminant analysis with respect to the goal of minimizing Bayes risks given multivariate distributions of factors. Based on this model, numerical analyses also perform with the application of a simulated dataset and an empirically observed dataset of traffic accidents in road segments. These examples essentially illustrate the process of Bayes risk minimization on predicating the safety status of road segments toward the objective of smallest misclassification rate. The paper finally concludes with a discussion of this methodology and several important avenues for future studies are also provided.

1. Introduction

Forecasting safety status of certain transportation facilities is crucial in the process of safety planning as well as black spot diagnoses [1]. It has gained great attention in the past safety forecasting models, which largely rely on the methods adopting regression analyses with respect to particular relationships between accident counts and several important factors. For example, earlier approaches used Poisson or negative binomial regression models [2, 3] to consider accident frequencies as count data. Further, advanced models are also developed with the additional concerns of statistical features such as zero-inflated framework [3], nonparametric specifications [4, 5], and multivariate distribution of responses [6, 7]. These models in fact attempted to accurately replicate the distributional property for aggregated crash data. However, these models highly depend on regression analyses, which assume the forecasting factors are fixed and therefore the

multicollinearity among factors is difficult to be addressed. Because factors are randomly distributed in nature and their stochastic relationships could provide additional information for enhancing the performance of forecasting models, it is necessary to involve the multivariate distributional characteristics of factors in addition to the conventional methods. To this end, this study suggests a mathematical approach based on the conception of Bayes risk minimization [8–10] to predict the safety status of transportation facilities with the consideration of the multicollinearity characteristics of factors.

The general idea of the proposed approach is to minimize the misclassification rate instead of maximizing the likelihood functions of regression models. In general forecasting models of road safety status, there are certain probabilities associated with the results that a road segment with high accident risks is classified as a safe road segment while a safer road could be also identified as a road with high accident

risks. The corresponding misclassifications are defined as type I and type II error, respectively, in statistical theory. As a result, the mathematical objective would be to look for a smallest rate for a weighted combination of the two types of errors. Such problem is indeed a statistical decision classification with a discrete action space and the action can be viewed as the categorization of safety status, and minimizing the risk for the decision problem will provide an optimum partition scheme for prediction purposes. Further, if a prior distribution of the market share of different categories of safety status is considered, then it will form a Bayes risk minimization problem.

The aforementioned method is a type of risk analysis [11] and also the foundation of discriminant analysis [12]. It has been applied in various areas in engineering. Specifically, it can be applied in the field of face recognition [13, 14], chemometrics [15], and even fluid mechanics [16]. Yet, such approach has not been applied in the area for analyzing transportation safety status and hence it is valuable to introducing this mathematical framework for research in transportation safety engineering. In addition, the introduced approach has advantages in terms of the asymptotic relative efficiencies comparing to logistic regression models [17] if the distributional assumptions of multicollinearity hold.

As a result, this study seeks to introduce a discriminant analysis framework based on the approach of Bayes risk minimization for forecasting the safety status of transportation facilities. Firstly, the mathematical formation of the proposed method is presented in the following section. Secondly, based on the model, numerical examples are illustrated according to a simulated dataset and an empirical observed dataset. Finally, a conclusion is provided with the summary of this study and several directions for future research are also provided.

2. The Method of Bayes Risk Minimization

A general forecasting problem is in the case that an individual or an outcome is classified into one of the k possible categories which are denoted by a vector $G = (g_1, g_2, \dots, g_k)^T$, and the set of factors are denoted by a vector $\mathbf{x} = (x_1, x_2, \dots, x_p)^T$. The classification regions for categories g_i are denoted by \mathcal{R}_i , respectively. Then, it assumes that the collection of all classification regions will form a partition of sample space of \mathbf{x} . Therefore, a natural classification rule can be expressed in the following: if $\mathbf{x} \in \mathcal{R}_i$, then the corresponding observation belongs to population g_i and otherwise it belongs to other categories. In the following analysis, $L(i, j)$ denotes the loss function for classifying an observation which belongs to category i into category j . The associated probability of such classification is $P(i, j)$ whose mathematical definition is presented in

$$P(i, j) = P(\mathbf{X} \in \mathcal{R}_j | G = g_i) = \int_{\mathbf{x} \in \mathcal{R}_j} f_i(\mathbf{x}) d\mathbf{x}, \quad (1)$$

where $f_i(\mathbf{x})$ is the probability density function for population i and the integration part represents the probability that an observation in population i is classified into population

j . For classification region, it is actually the sample space partitioned in which each partition will correspond to categories of the subject. As aforementioned, the optimum region $\mathcal{R}_i, i = 1, \dots, k$ will form a partition of the sample space such that the corresponding Bayes risk $r(\pi, R)$ is minimized.

The definition of Bayes risks for this problem is presented in

$$r(\pi, \mathcal{R}) = \sum_{i=1}^k \sum_{j=1}^k [\pi_i P(i, j) L(i, j)]. \quad (2)$$

In the above definition, Bayes risk actually measures the expected losses resulting from misclassifications, and π_i is the market sharing of each category. The idea of adopting such prior information is to weight the expected losses from coming observations by consideration of their overall proportion in each category. Therefore, the partition criterion will try to avoid errors from misclassification of populations which have a large market share. Because minimizing the Bayes risk is equivalent to minimizing the posterior risk at every point of \mathbf{x} , the optimum partition boundaries are required to separate the sample space such that, for each sample point, the error rate reaches the minima. According to Bayes theorem, the posterior distribution of an observation coming from g_i is given in

$$P(G = g_i | \mathbf{x}) = \frac{\pi_i f_i(\mathbf{x})}{\sum_{r=1}^k [\pi_r f_r(\mathbf{x})]}. \quad (3)$$

Accordingly the posterior risk is defined as the expected conditional loss and it is expressed in (4). To be noted, the posterior risk depends on specific value of j which is the category that observation \mathbf{x} will be classified into according to the partition rules. Consider

$$R(\mathbf{x}, j) = \sum_{i=1}^k [P(G = g_i | \mathbf{x}) L(i, j)] = \sum_{i=1}^k \left[\frac{\pi_i f_i(\mathbf{x}) L(i, j)}{\sum_{r=1}^k \pi_r f_r(\mathbf{x})} \right]. \quad (4)$$

Therefore, for each observation \mathbf{x} , the decision rule is characterized by selecting j such that the above posterior risk is the smallest. If $L(i, j)$ is defined as the regular 0-1 loss in

$$L(i, j) = \begin{cases} 0, & \text{if } i = j \\ 1, & \text{if } i \neq j, \end{cases} \quad (5)$$

then the posterior risk will be simplified to the form in

$$R(\mathbf{x}, j) = \sum_{i \neq j} P(G = g_i | \mathbf{x}) = 1 - P(G = g_j | \mathbf{x}). \quad (6)$$

Hence, the problem is equivalent to selecting j such that $P(G = g_j | \mathbf{x})$ is maximized. In other words, the best classification rule will select the population that has the largest conditional probabilities given the observed value of \mathbf{x} . According to the general mathematic formulation, it is ready to analyze the problem given the distribution of the factor vector or given the distribution family and leave the parameters to be estimated. In this section, cases that the

measurement vector is distributed as multivariate normal with homogeneous variance-covariance matrix across different categories are assumed. As it will be illustrated shortly, the resulting partition boundary is linear in \mathbf{x} and such scenario is also called linear discriminant analysis [18]. For a problem of k categories, it supposes that the vector \mathbf{X} is distributed as multivariate normal distribution $N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$ in category g_i for all i . The probability density function of \mathbf{X} in category g_i is given by

$$f(\mathbf{x} | G = g_i) = \frac{1}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_i)\right]. \quad (7)$$

Based on Bayes theorem, the posterior $P(G | \mathbf{x})$ is given by

$$P(G = g_j | \mathbf{x}) = \frac{f(\mathbf{x} | G = g_j) P(G = g_j)}{\sum_{i=1}^k f(\mathbf{x} | G = g_i) P(G = g_i)} = \frac{f(\mathbf{x} | G = g_j) \pi_j}{\sum_{i=1}^k f(\mathbf{x} | G = g_i) \pi_i}. \quad (8)$$

For 0-1 loss function, the decision can be made by choosing the largest $P(G = g_j | \mathbf{x})$ at each \mathbf{x} , and the partition of sample space can be formulated through all the pairwise classification boundaries. Specifically, the classification boundaries between populations i and j can be described by balancing their odds in

$$\begin{aligned} \frac{P(G = g_i | \mathbf{x})}{P(G = g_j | \mathbf{x})} &= \frac{f(\mathbf{x} | G = g_i) \pi_i}{f(\mathbf{x} | G = g_j) \pi_j} \\ &= \frac{\pi_i}{\pi_j} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) + \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_j)^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_j)\right] = 1 \quad (9) \\ &= \frac{\pi_i}{\pi_j} \exp\left[\mathbf{x}^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) - \frac{1}{2}(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)\right] = 1 \end{aligned}$$

or expressed by the log odds in (10) for convenience:

$$\begin{aligned} \log \frac{P(G = g_i | \mathbf{x})}{P(G = g_j | \mathbf{x})} &= \mathbf{x}^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) \\ &\quad - \frac{1}{2}(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) \quad (10) \\ &\quad + \log \frac{\pi_i}{\pi_j} = 0. \end{aligned}$$

The first term on the right-hand side of (10) is called discrimination function [19]. In order to minimize the Bayes risk, the classification boundary between g_i and g_j is given by

$$\mathbf{x}^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) - \frac{1}{2}(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) + \log \frac{\pi_i}{\pi_j} = 0. \quad (11)$$

Since the discriminant function [20] is linear in \mathbf{x} , the pairwise decision boundary is a hyperplane in \mathbb{R}^P . Even though there are $k(k-1)/2$ pairwise classifications corresponding to the combination of i and j , only $k-1$ pairwise classifications are irredundant. In practice, applications can use the maximum likelihood estimates [21] for parameters π_i , $\boldsymbol{\mu}_i$, and $\boldsymbol{\Sigma}$. Therefore, π_i is estimated by its group proportion $\hat{\pi}_i = n_i/n$ and

$$\hat{\boldsymbol{\mu}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{x}_{ij} \quad (12)$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} (\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)(\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)^T.$$

Substituting these estimates, the estimated classification boundary based on their log odds is

$$\mathbf{x}^T \hat{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_j) - \frac{1}{2}(\hat{\boldsymbol{\mu}}_i + \hat{\boldsymbol{\mu}}_j)^T \hat{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_j) + \log \frac{\hat{\pi}_i}{\hat{\pi}_j} = 0. \quad (13)$$

If two categories are considered, the above analysis reduces to the two population problems and the decision boundary is readily formed by a single hyperplane. For the two population cases, the estimated discriminant function is also proportional to the famous Fisher discriminant function [19]. To be noted, Bayes risk minimization is in fact a different meaning other than the accident risk, which scales the risk of being in an incorrect estimation of parameters.

3. Numerical Examples

With the above specified mathematical formulation for the forecasting problems under minimized Bayes risks, the analysis in this section first presents a numerical example on simulated samples and then on an empirical observed dataset. In order to illustrate the estimated partition of sample space, a bivariate sample space is adopted and hence bivariate normal distributions are assumed for two groups of populations within different categories individually. Equation (14) gives predefined parameters for simulation purposes, namely, the mean vectors and variance-covariance matrix for the two groups of populations:

$$\boldsymbol{\mu}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\mu}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.3 \\ 0.3 & 1 \end{pmatrix}. \quad (14)$$

To be noted, the mean parameters for the two groups of populations choose different values whereas this study

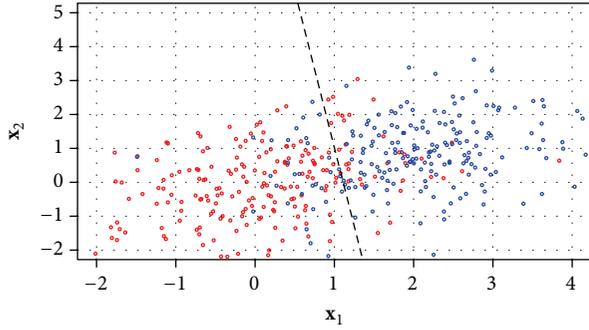


FIGURE 1: Discriminant partitions of sample space for simulated dataset.

uses a homogeneous variance-covariance matrix for the two populations. Under this assumption, the partition boundary will be a straight line [22] that separates the sample space and also provide prediction criterion for future cases. However, if the equal covariance assumption fails [22], the pairwise log odds will be a quadratic function of the measurement vector \mathbf{x} . Therefore, the decision boundary will be quadratic in \mathbf{x} and the corresponding analysis is then called quadratic discriminant analysis. This study is a preliminary step in this field and methods under loose assumptions of the covariance matrix are left as a direction for future research.

With such specified parameters, the dataset is simulated with 200 samples for each group of population. Figure 1 illustrates the scatter plot of the two datasets, in which they are represented by different colors and the dashed line is the estimated discriminant boundary which partitions the sample space into two parts under the minimized Bayes risks.

Because the two populations are randomly distributed with overlapped domain, their samples are also overlapped with certain probabilities. It is impossible to find a partition to perfectly separate the two populations and hence misclassification always occurs due to the overlapping. Because the illustrated examples use two populations with different mean vectors, they will concentrate on different regions around their mean points, which will provide effective information to make future prediction. If the mean points are far from each other, the classification error will be small and vice versa. Thus, for two populations with adjacent mean points, any partition may not be able to reduce error rate due to the fact that they actually lack distinct features for identification. Consequently, for empirical analyses, it is important to choose features that can provide effective information for distinguishing different populations.

Then the following section presents an empirical analysis based on observed crash dataset for road segments. The data is collected in Pikes Peak Area, Colorado, USA, for the period from July 2006 to December 2010. In this data, the forecasting subject will be the safety status of road segments, which can be simply categorized as safe road or unsafe road to reflect the level of accident risks. In general, the accident risk measures the chance of being in accidents or the magnitude of encountered harmfulness from accidents. For example, crash count, crash rates, and equivalent property damage

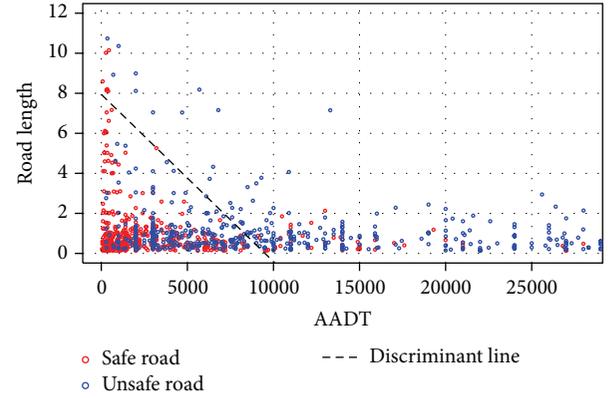


FIGURE 2: Discriminant partitions of sample space for observed crash dataset.

only crashes are all risk representations. In this study, the occurrence of at least one accident is adopted as the risk of accidents to account for the safety status of road segments. For such subject, past experience has identified the fact that the variables, annual average daily traffic (AADT), and the length of the road segment may show significant impacts on the occurrence of accidents. Thus, this study will adopt these factors to make predictions on the safety status of road segments. Consider

$$\begin{aligned} \hat{\boldsymbol{\mu}}_1 &= \begin{pmatrix} 3.4 \times 10^3 \\ 1.0 \end{pmatrix}, & \hat{\boldsymbol{\mu}}_2 &= \begin{pmatrix} 12.0 \times 10^3 \\ 1.1 \end{pmatrix}, \\ \hat{\boldsymbol{\Sigma}} &= \begin{pmatrix} 9.6 & -3.3 \\ -3.3 & 3.8 \end{pmatrix} \times 10^7. \end{aligned} \quad (15)$$

Equation (15) gives the estimated parameters from the dataset and it is computed from maximized likelihood estimation. Here, $\hat{\boldsymbol{\mu}}_1$ and $\hat{\boldsymbol{\mu}}_2$ are in fact the mean vectors for safe group and unsafe group, respectively. For a specific mean vector, the first element is corresponding to AADT and the second element is corresponding to segment length. Since the true parameters are unknown, the following analysis will be based on the estimated parameters for both datasets of safe road segments and unsafe road segments. As usual, the estimated mean points are allowed to be different while the variance-covariance matrix is still assumed to be the same. Moreover, the mean point is estimated based on the individual dataset of each population while, during the estimation of the variance-covariance matrix, all samples are used. Equation (12) provides the mathematical formulation for estimating these parameters.

For numerical computation, the only concern about this problem is to make sure that the variance-covariance matrix is invertible, which is usual and easy to satisfy. Figure 2 illustrates the partition as well as all the locations of samples of safe roads and unsafe roads. Even though the partition line can provide a solution to classify a road segment into one particular category of safety status, it may not be optimum due to the fact that the assumption of bivariate normal distributions for the entire population placed before the

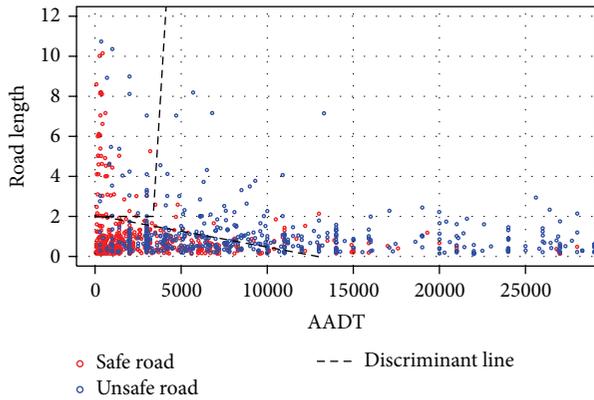


FIGURE 3: Improved partitions of sample space for crash dataset.

analysis may not hold. Since the analysis is based on empirical observed data, the multivariate distribution among factors may not be able to highly match the initial assumptions. Thus, it is necessary to improve current partition scheme with respect to the consideration of more realistic distributional assumptions. By observation of these sample points, for the roads with smaller segment length, the population distribution behaves differently with the roads which have relatively larger segment length, since the former presents a different shape of distribution form of sample points which are concentrated at different mean point.

Therefore, it is necessary to consider this phenomenon and try to find methods to overcome this issue. This analysis accordingly proposed a strategy in which samples with different behaviors are treated individually for partitioning purposes. The following analysis will conduct two independent partition analyses according to the samples with road length greater than two miles and the rest of samples with road length less than two miles. Figure 3 illustrates the improved partition line in the entire sample spaces based on the results from two individual classification problems. The broken line actually functions as a boundary to classify a future observation. There might also be other solutions to deal with this issue. For example, a better method could optimize the separation of populations with the concern of different behaviors of distribution and this type of methods leaves as a direction for future research.

4. Discussion and Conclusions

The mathematical approaches based on the minimization of Bayes risk are effective to classify an unknown outcome into one of several possible categories with respect to the smallest misclassification rate. It is a widely used technique which has been applied in various areas. However it is rarely mentioned in the past studies of forecasting the safety status of transportation facilities, which largely rely on the regression framework. Because of the nature of regression that the explanatory variables are given without the consideration of their stochastic behaviors, the multicollinearity among factors is less considered. The occurrence of accident

is the result of complicate processes controlled by many influence factors including the road geometric design, traffic flow characteristics, and characteristics of drivers. In fact, these factors are highly correlated in nature. Therefore it is necessary and secure to assume that multicollinearity of factors exists in analyses. Under this circumstance, this study seeks to introduce a discriminant analysis framework as a tool to classify transportation facilities in categories defined by different level of safety. Instead of maximizing the conditional likelihood of responses, the introduced method in fact attempts to minimize the Bayes risks for the classification problem. Under such specification, the current study illustrates numerical examples on a simulated dataset as well as an empirical observed dataset. It finally shows that this approach also constructs a flexible framework for conducting the analysis on prediction of traffic safety status in the field of transportation engineering. In addition, as a preliminary research into this field, next section will suggest several directions for future studies.

First, it is necessary to analyze different loss functions in terms of nonuniform loss weight to different types of errors with the consideration of the tradeoff between the additional accident loss introduced by misclassifying an unsafe road into a safe road and the cost for improving an originally safe road but misclassified into the unsafe category. Second, an inhomogeneous variance-covariance matrix may exist among different group of populations with different status of safety and hence future studies may take this into account. And, with more flexible parameters and looser assumptions, the partition may provide more realistic and accurate classification boundaries. Third, it is true that the actual distribution of factors might be reflected better by other types of multivariate formulation. Inspection on any other multivariate distributions even nonparametric specifications is necessary for future studies. Finally, it is interesting to compare the prediction results to the traditional regression analysis with simulated datasets or empirical observed datasets to further understand the statistical characteristics for transportation safety datasets.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (nos. 51208032 and 71210001) and Fundamental Research Funds for the Central Universities (no. 2013JBM008).

References

- [1] A. Gregoriades and K. C. Mouskos, "Black spots identification through a Bayesian Networks quantification of accident risk index," *Transportation Research C: Emerging Technologies*, vol. 28, pp. 28–43, 2013.

- [2] A. Tarko and M. Tracz, "Accident prediction models for signalized crosswalks," *Safety Science*, vol. 19, no. 2-3, pp. 109-118, 1995.
- [3] S.-P. Miaou, "The relationship between truck accidents and geometric design of road sections: poisson versus negative binomial regressions," *Accident Analysis & Prevention*, vol. 26, no. 4, pp. 471-482, 1994.
- [4] Y. Xie and Y. Zhang, "Crash frequency analysis with generalized additive models," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2061, pp. 39-45, 2008.
- [5] L. Ma and X. Yan, "Examining the nonparametric effect of drivers' age in rear-end accidents through an additive logistic regression model," *Accident Analysis and Prevention*, vol. 67, pp. 129-136, 2014.
- [6] J. Ma and K. M. Kockelman, "Bayesian multivariate poisson regression for models of injury count, by severity," *Transportation Research Record*, vol. 1950, pp. 24-34, 2006.
- [7] Y.-C. Chiou and C. Fu, "Modeling crash frequency and severity using multinomial-generalized Poisson model with error components," *Accident Analysis and Prevention*, vol. 50, pp. 73-82, 2013.
- [8] V. Goel and W. J. Byrne, "Minimum Bayes-risk automatic speech recognition," *Computer Speech and Language*, vol. 14, no. 2, pp. 115-135, 2000.
- [9] U. K. Müller, "Risk of Bayesian inference in misspecified models, and the sandwich covariance matrix," *Econometrica*, vol. 81, no. 5, pp. 1805-1849, 2013.
- [10] F. Ruggeri, "On some optimal Bayesian nonparametric rules for estimating distribution functions," *Econometric Reviews*, vol. 33, no. 1-4, pp. 289-304, 2014.
- [11] Y. Yan and B. Suo, "Risks analysis of logistics financial business based on evidential bayesian network," *Mathematical Problems in Engineering*, vol. 2013, Article ID 785218, 8 pages, 2013.
- [12] Y. Rahulamathavan, R. C.-W. Phan, J. A. Chambers, and D. J. Parish, "Facial expression recognition in the encrypted domain based on local fisher discriminant analysis," *IEEE Transactions on Affective Computing*, vol. 4, no. 1, pp. 83-92, 2013.
- [13] J. Chai, H. Liu, and Z. Bao, "Generalized re-weighting local sampling mean discriminant analysis," *Pattern Recognition*, vol. 43, no. 10, pp. 3422-3432, 2010.
- [14] Z. Li, D. Lin, and X. Tang, "Nonparametric discriminant analysis for face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 4, pp. 755-761, 2009.
- [15] L. Clemmensen, T. Hastie, D. Witten, and B. Ersbøll, "Sparse discriminant analysis," *Technometrics*, vol. 53, no. 4, pp. 406-413, 2011.
- [16] C.-Y. Chen, "Using discriminant analysis to determine the breaking criterion for an ISW propagating over a ridge," *Environmental Fluid Mechanics*, vol. 10, no. 5, pp. 577-586, 2010.
- [17] B. Efron, "The efficiency of logistic regression compared to normal discriminant analysis," *Journal of the American Statistical Association*, vol. 70, no. 352, pp. 892-898, 1975.
- [18] H. Wakaki, "Comparison of linear and quadratic discriminant functions," *Biometrika*, vol. 77, no. 1, pp. 227-229, 1990.
- [19] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, NY, USA, 2003.
- [20] J. Cornfield, "Discriminant functions," *Review of the International Statistical Institute*, vol. 35, pp. 142-153, 1967.
- [21] J. H. Friedman, "Regularized discriminant analysis," *Journal of the American Statistical Association*, vol. 84, no. 405, pp. 165-175, 1989.
- [22] S. Marks and O. J. Dunn, "Discriminant functions when covariance matrices are unequal," *Journal of the American Statistical Association*, vol. 69, pp. 555-559, 1974.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

